Analysis of social mobility using the M index

Ben Jann and Simon Seiler

Institute of Sociology, University of Bern
ben.jann@soz.unibe.ch, simon.seiler@icer.unibe.ch

Workshop at Universidad de Sevilla
November 16, 2019

Motivation

- Sociological research on intergenerational mobility faces specific methodological challenges due to the categorical nature of one of its key variables: class position.
- Much effort has been put into designing methods that are "margin-free". These methods try to quantify the "social fluidity" of a society net of structural mobility.
- We argue that this quest has gone too far – at least for research topics such as the international comparison of social mobility – and we propose an alternative methodology.

Outline

1. How to quantify "social fluidity"?
2. The M-index
3. Example 1: Reanalysis of Long and Ferrie (2013)
4. Example 2: Social mobility in 20th century Switzerland
5. Conclusions

Absolute mobility

- Societies in which social origin has little influence on an individual's social position are said to be "open" or "socially fluid".
- Social fluidity often goes along with relatively high degrees of *absolute* mobility – in the sense that the observed social position of a person differs from the social position of her or his family of origin.
- Early mobility research focused on absolute mobility, but absolute mobility may not always be a good measure of social fluidity.

Structural mobility

- High absolute mobility is neither necessary nor sufficient for social fluidity.
- For example, absolute mobility may be high despite strong effects of social origin, because the class structure changes from one generation to the next.
  - During rapid industrialization, the working class grows while non-industrial classes shrink. In such a situation, many descendants of a non-industrial class will be "forced" to be socially mobile, because there are not enough non-industrial positions within the class structure of their own generation.
- Such forced mobility is often labeled "structural mobility". When determining the social fluidity of a society, a distinction has to be made between structural mobility and other forms of mobility.

Relative mobility and log-linear models

- The mobility that is relevant for social fluidity is called *relative* mobility (or exchange mobility) .
- Relative mobility is a comparative concept that has to do with inequality of opportunities: relative social mobility is high if the chances of attaining a certain position are similar for all social origins. In this case, social origin is inconsequential and the society can be said to be open or socially fluid.
- Relative mobility can be formalized by means of odds ratios (odds ratio of a person from origin i compared to a person from origin j of attaining position k instead of l) or, equivalently, by so-called log-linear models.

Comparing societies using the unidiff model

- Patterns of odds ratios or parameters of log-linear models do not provide an easy-to-interpret overall fluidity measure. This makes comparing results over time or between countries difficult.
- A popular approach to solve this problem is the so-called "unidiff" model by Erikson and Goldthorpe (1992) a.k.a. the "log multiplicative layer effects model" by Xie (1992).
- The unidiff model is a log-linear model in which a distinction is made between the association pattern between origin and destination, and the "strength" of these associations: While the pattern is common to all compared societies, it is allowed to vary uniformly in strength between them.

The unidiff model

- Think of a three-way table:

$$\text{destination } (Y) \times \text{origin } (X) \times \text{cohort } (Z)$$

- The log of the cell frequencies μ_{jkl} can be described using a saturated log-linear model:

$$\ln(\mu_{jkl}) = \lambda + \lambda_j^X + \lambda_k^Y + \lambda_l^Z + \lambda_{jk}^{XY} + \lambda_{jl}^{XZ} + \lambda_{kl}^{YZ} + \lambda_{jkl}^{XYZ}$$

- Such a model perfectly captures the data, but is not very useful for interpretation due to the large number of parameters.
- We can simplify the model to the so-called constant fluidity model that assumes the relation between destination and origin to be constant across cohorts:

$$\ln(\mu_{jkl}) = \lambda + \lambda_j^X + \lambda_k^Y + \lambda_l^Z + \lambda_{jk}^{XY} + \lambda_{jl}^{XZ} + \lambda_{kl}^{YZ}$$

- This is also not very useful because we are interested in analyzing change over cohorts and don't want to assume these changes away.

The unidiff model

- The idea of the unidiff model now is to assume that the general association pattern between Y and X says the same, but that the pattern can be more or less pronounced.
- To achieve this, λ_{jk}^{XY} is replaced by $\phi_l \psi_{jk}^{XY}$ where ψ_{ij}^{XY} describes the common pattern and ϕ_l is a cohort-specific scaling factor:

$$\ln(\mu_{jkl}) = \lambda + \lambda_j^X + \lambda_k^Y + \lambda_l^Z + \lambda_{jl}^{XZ} + \lambda_{kl}^{YZ} + \phi_l \psi_{jk}^{XY}$$

- Note that the unidiff model can also be expressed at the level of individual observations instead of cell frequencies. In this case, the model is very similar to a multinomial logit model:

$$\Pr(Y = k | X_i, Z_i) = \frac{\exp(\alpha_k + Z_i \beta_k + Z_i \phi X_i \psi_k)}{\sum_{h=1}^{J} \exp(\alpha_h + Z_i \beta_h + Z_i \phi X_i \psi_h)}$$

where X_i and Z_i are vectors of indicator variables.

Bringing the margins back in

- As long as the uniformity assumption holds (which can be tested), the "strength" parameters of the unidiff model seem like an elegant and parsimonious way to evaluate overall social fluidity.
- So what might be the problem?
- The unidiff model captures the pattern and strength of class barriers, but each barrier receives the same weight irrespective of the proportion of the society that faces the barrier.

Bringing the margins back in

- Though experiment:
  - Think of a society in which social origin effects are limited to a single class, say, farmers. For all other classes, odds ratios are equal to one.
  - Assume that none of the class barriers change over time.
  - Assume that the proportion of the selective class declines over time and, eventually, becomes zero.
  - In the unidiff model, this society will always have the same level of fluidity as long as the selective class exists. Once the class vanishes, however, the society will be described as completely open.
- Thus, if we apply the aggregation rule built into the unidiff model for making substantive generalizations from individual class barriers to the overall openness of a society, we accept that a large class makes the same contribution to a society's social rigidity as a class that almost disappeared – while a class that has completely disappeared contributes nothing.

Bringing the margins back in

- Some desired properties of a measure of social fluidity
  - Should be zero in case of complete independence of origin and destination.
  - Should be margin-free in the sense that an existing association between origin and destination is identified net of structural mobility (i.e., should be based on a comparison of conditional distributions).
  - However, should take into account the marginal distribution as far as it affects the relevance of existing dependencies between origin and destination due to the proportion of society affected by these associations.
  - Differences in the measure should be decomposable into a part due to differences in margins and a part due to differences in associations.
  - The measure should be decomposable by subgroups (e.g. different destinations or origins).

1 How to quantify "social fluidity"?

2 The M-index

3 Example 1: Reanalysis of Long and Ferrie (2013)

4 Example 2: Social mobility in 20th century Switzerland

5 Conclusions

The M-index

- A measure that meets these properties can be derived from information theory (Theil 1970 and Theil and Finizza 1971; also see Mora and Ruiz-Castillo 2011).
- To quantify the effect of social origin we ask: How much can we learn, on average, about a person's class by knowing the person's social origin?
- The question is answered by computing the difference between the *a priori* (not knowing the origin) and *a posteriori* (knowing the origin) information gain of actually observing the person's class.

The M-index

- A priori entropy of Y (destination class):

$$T(P_Y) = -\sum_{k=1}^{K} p(y_k)\ln(p(y_k))$$

- A posteriori entropy of Y, given X (origin class):

$$T(P_{Y|X}) = -\sum_{j=1}^{J} p(x_j)\sum_{k=1}^{K} p(y_k|x_j)\ln(p(y_k|x_j))$$

- M-index:

$$M = T(P_Y) - T(P_{Y|X}) = \sum_{j=1}^{J}\sum_{k=1}^{K} p(y_k, x_j)\ln\left(\frac{p(y_k|x_j)}{p(y_k)}\right)$$

The M-index

- The M-index can also be expressed as a simple average of observation-specific components:

$$M = E(m_i) = \frac{1}{N}\sum_{i=1}^{N} m_i = \frac{1}{N}\sum_{i=1}^{N} \ln\left(\frac{\Pr(Y_i|X_i)}{\Pr(Y_i)}\right)$$

- From this formulation we see that ...
  - M can easily be decomposed by subgroups.
  - M is easy to compute, even if X is multidimensional.
  - additional information (V) can easily be integrated ("partial" M).
    - a priori model: $\Pr(Y_i|V_i)$
    - a posteriori model: $\Pr(Y_i|V_i, X_i)$

Estimation

1. Estimate a multinomial logit (or whatever is appropriate) of Y on V (where V may be empty) and predict $\Pr(Y_i|V_i)$.
2. Estimate a multinomial logit of Y on V and X and predict $\Pr(Y_i|V_i, X_i)$.
3. Compute m_i for each observation

$$\widehat{m}_i = \ln\left(\frac{\widehat{\Pr}(Y_i|V_i, X_i)}{\widehat{\Pr}(Y_i|V_i)}\right)$$

4. Analyze m_i, by taking (subgroup) averages or running regressions including whatever covariates seem appropriate for the research question at hand.

Estimation

- If M is to be compared between countries or across birth cohorts (or, more generally, by some set of variables Z), we can repeat steps 1–3 individually or use joint models with appropriate interaction terms (we follow the latter approach; this allows, for example, using continuous Z variables).
- Packing all involved equations together into a GMM system is a way to obtain consistent standard errors. Another approach is the bootstrap. In many cases, however, we observe that standard analysis of m_i ignoring the fact that it is based on estimated quantities provides a good approximation.

Counterfactual decomposition

- When comparing M between countries or birth cohorts, the part of the difference due to differences in the margins can be identified using counterfactual decompositions. The basic idea is to compute M based on simulated data where the associations are maintained but the marginal distributions are exchanged.
- Silber and Spadaro (2011) suggest a procedure based on “raking”, which operates at the level of cell frequencies in a two-way table.
  - The procedure can be be generalized to situations in which multinomial logit models are used to estimate the outcome probabilities.
- In the context of the analysis of segregation, Mora and Ruiz-Castillo (2009) propose an alternative decomposition (also see Di Prete et al. 2017).

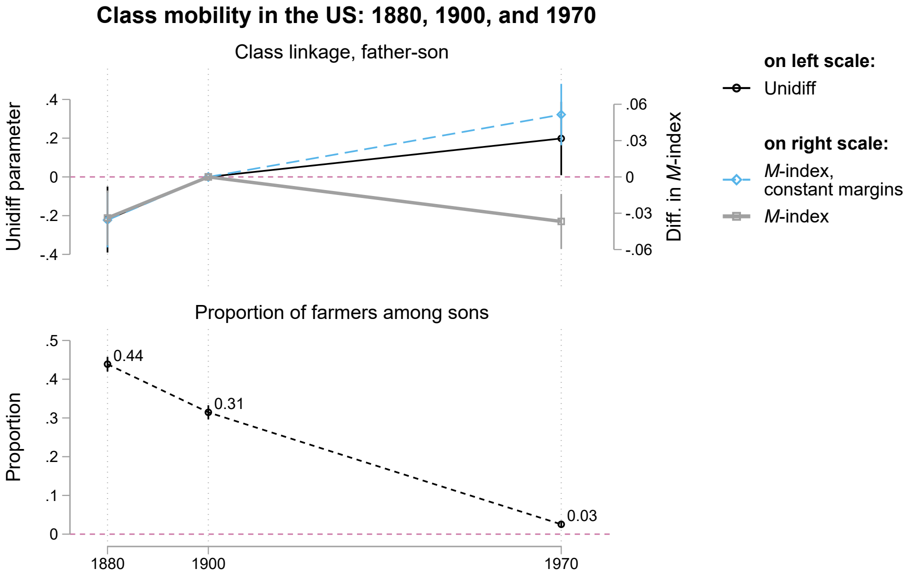
1. How to quantify “social fluidity”?
2. The M-index
3. Example 1: Reanalysis of Long and Ferrie (2013)
4. Example 2: Social mobility in 20th century Switzerland
5. Conclusions

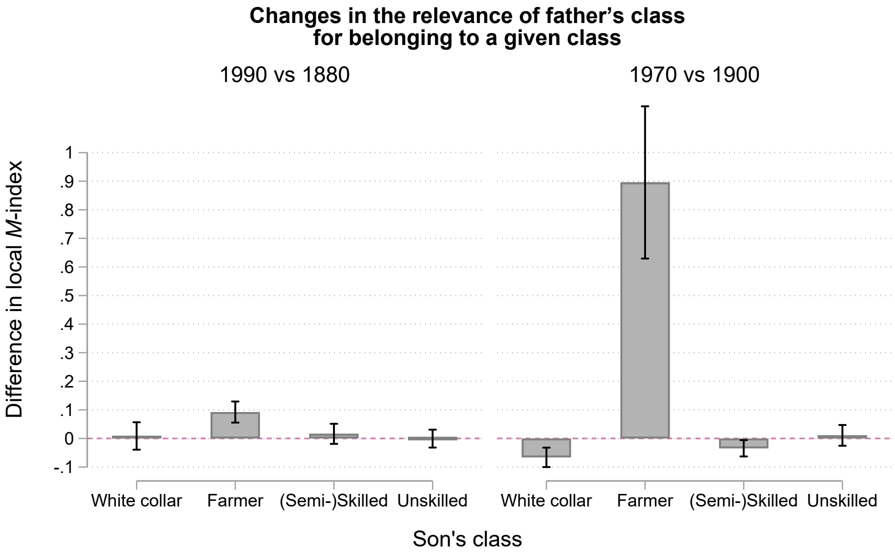
Reanalysis of Long and Ferrie (2013)

- Long and Ferrie (2013) analyzed social mobility in Great Britain and the United States after 1850.
- Controversial conclusion: the US was more open in the 19th century than in the 20th century (finding based on both their own measure and the unidiff model).
- Our reanalysis:
  - We can confirm their conclusion using an unidiff model: Class barriers became more rigid from 1880 to 1900 and again from 1900 to 1970.
  - The M-index leads to the same conclusion – if we adjust the margins in 1880 and 1970 to be the same as in 1900 (decomposition by Silber and Spadaro 2011).
  - However, using the raw M-index, results look very different. We now have an increase in social mobility between 1900 and 1970.

Reanalysis of Long and Ferrie (2013)

- Why is there such a difference between unidiff and the *M*-index?
- As Hout and Guest (2013) as well as Xie and Killewald (2013) rightly criticized, the unidiff results are driven by the (increasingly) strong rate of farmers recruited among sons of farmers, while at the same time the proportion of farmers decreased dramatically.
- The *M*-index takes the decline of the proportion of farmers into account, whereas the unidiff model does not.
- The strong increase in class linkage for farmers can be seen nicely when disaggregating the change in the *M*-index by class.


Class mobility in the US: 1880, 1900, and 1970
Class linkage, father-son
Unidiff parameter
.4
.2
0
-.2
-.4
.06
.03
0
-.03
-.06
Diff. in M-index
on left scale:
Unidiff
on right scale:
M-index,
constant margins
M-index
Proportion of farmers among sons
Proportion
.5
.4
.3
.2
.1
0
0.44
0.31
0.03
1880
1900
1970


Changes in the relevance of father's class
for belonging to a given class
1990 vs 1880
1970 vs 1900
Difference in local M-index
1
.9
.8
.7
.6
.5
.4
.3
.2
.1
0
-.1
White collar
Farmer
(Semi-)Skilled
Unskilled
White collar
Farmer
(Semi-)Skilled
Unskilled
Son's class

1 How to quantify "social fluidity"?

2 The M-index

3 Example 1: Reanalysis of Long and Ferrie (2013)

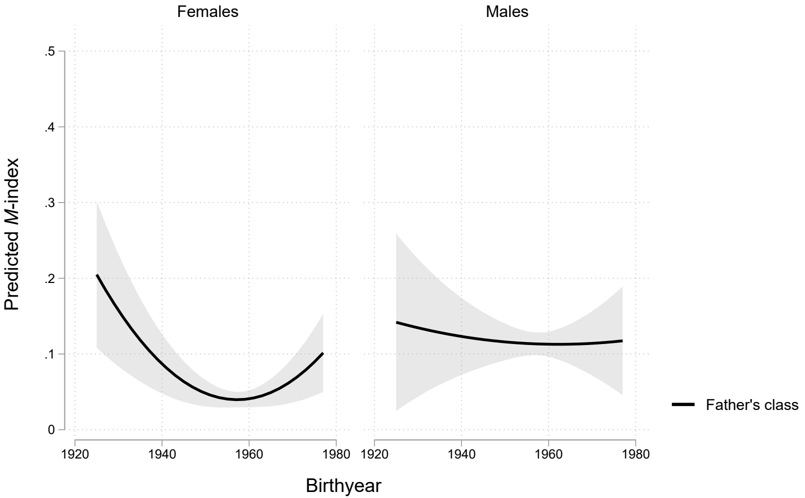
4 Example 2: Social mobility in 20th century Switzerland

5 Conclusions

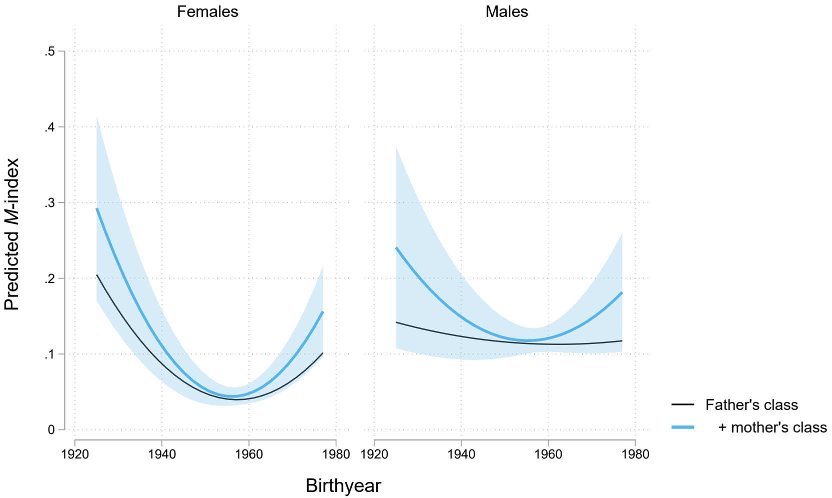
Social mobility in 20th century Switzerland

- Our second example illustrates that the *M*-index makes it easy to include multiple origin variables into the analysis, and that this may matter for the results.
- Data from Jann and Seiler (2014); includes 20 Swiss surveys with a total of about 24,000 respondents; we analyze (slightly simplified) EGP classes (Erikson et al. 1983) of respondents aged 35-69 (including "homemaker" as a separate class)
- We model social mobility over birth cohorts, separately for men and women, using different sets of origin variables (assuming additive effects):
  1. Father's class
  2. Father's class and mother's class
  3. Father's class and education, mother's class and education (educational categories: low, middle, high)

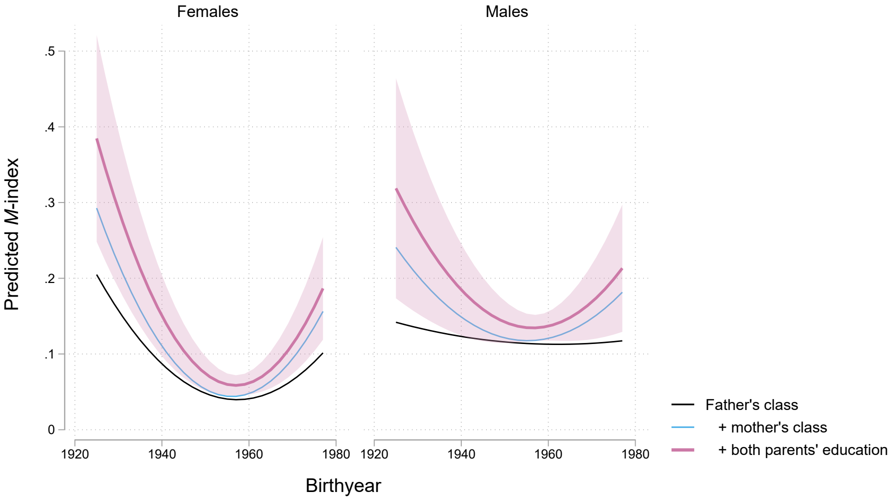
Father's class only



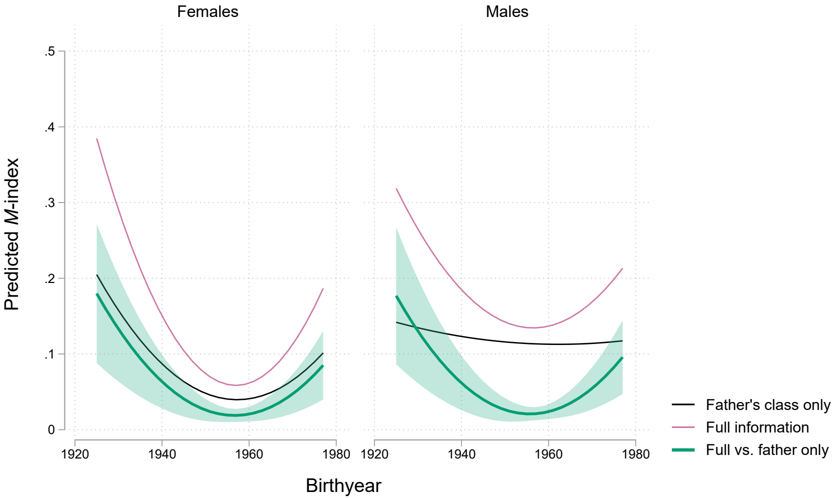
Class of both parents



Class and education of both parents



Difference full model vs. father's class only



Social mobility in 20th century Switzerland

- Results show that mothers matter.
- In particular for men, considering information on mothers radically changes our conclusions. When only considering father's occupational class, there is hardly any trend in origin effects; however, if taking into account both parents' class, we find a U-shaped trend (as for women).
- It thus seems that, in Switzerland, social mobility increased until about birth cohorts of 1960, but then social origin effects started to increase again. Ignoring information on mothers, we would have missed that this pattern is true for both men and women.

1 How to quantify "social fluidity"?

2 The M-index

3 Example 1: Reanalysis of Long and Ferrie (2013)

4 Example 2: Social mobility in 20th century Switzerland

5 Conclusions

Conclusions

- We believe that the *M*-index is a very fruitful and versatile approach to study social mobility in international comparison.
- Like log-linear/unidiff models, the *M*-index is margins-free in the sense that it is based on conditional distributions/odds ratios.
- However, the *M*-index brings the margins back in when aggregating these class barriers. It thus takes into account how "important" the different barriers are in a given society.
- The *M*-index offers great flexibility in terms of how the class barriers are modeled (multiple and possibly continuous origin variables, control variables such as survey dummies, sampling weights and survey design, . . .) and how the results are analyzed (disaggregation by classes, effects of aggregate-level variables, decomposition, . . .)
- Basic implementation is straight forward; no special software required (although we will provide some dedicated programs).

References I

- DiPrete, T.A., T. Bol, C.C. Eller, H.G. van de Werfhorst. 2017. School-to-Work Linkages in the United States, Germany, and France. American Journal of Sociology 122(6): 1869–1938.
- Erikson, R., J.H. Goldhorpe. 1992. The constant flux. A study of class mobility in industrial societies. Oxford: Clarendon.
- Erikson, R., J.H. Goldthorpe, L. Portocarero. 1983. Intergenerational Class Mobility and the Convergence Thesis: England, France and Sweden. The British Journal of Sociology 34(3): 303–343.
- Ganzeboom, H.B.G. 2013. ISCO-88 Codes for Parental Occupations in the European Social Survey, Rounds 1-2-3-4-5. Amsterdam: VU-University. Version 1 (July 18, 2013), available from `http://www.harryganzeboom.nl/ESS-DEVO/ess12345_fmisko.zip`.
- Hout, M., A.M. Guest. 2013. Intergenerational Occupational Mobility in Great Britain and the United States Since 1850: Comment. American Economic Review 103(5): 2021–2040.
- Kuitto, K. 2011. More than just Money: Patterns of Disaggregated Welfare Expenditure in the Enlarged Europe. Journal of European Social Policy 21(4): 348– 364.
- Jann, B., S. Seiler. 2014. A New Methodological Approach for Studying Intergenerational Mobility With an Application to Swiss Data. University of Bern Social Sciences Working Paper No. 5 (`http://ideas.repec.org/p/bss/wpaper/5.html`).

References II

- Long, J., J. Ferrie. 2013 . Intergenerational Occupational Mobility in Great Britain and the United States Since 1850. American Economic Review 103(4): 1109–1137.
- Mora, R., J. Ruiz-Castillo. 2009. The Invariance Properties of the Mutual Information Index of Multigroup Segregation. P. 33–53 in: Y. Flückiger, S.F. Reardon, J. Silber (eds.). Occupational and Residential Segregation. Bingley: Emerald.
- Mora, R., J. Ruiz-Castillo. 2011. Entropy-Based Segregation Indices. Sociological Methodology 41: 159–194.
- Silber, J., A. Spadaro. 2011. Inequality of Life Chances and the Measurement of Social Immobility. P. 129–154 in: M. Fleurbaey, M. Salles, J.A. Weymark (eds.). Social Ethics and Normative Economics. Berlin: Springer.
- Theil, H. 1970. On the estimation of relationships involving qualitative variables. American Journal of Sociology 76(1): 103–154.
- Theil, H., A.J. Finizza. 1971. A note on the measurement of racial integration of schools by means of informational concepts. Journal of Mathematical Sociology 1(2): 187–193.
- Xie, Y. 1992. The Log-Multiplicative Layer Effect Model for Comparing Mobility Tables. American Sociological Review 57(3): 380–395.
- Xie, Y., A. Killewald. (2013). Intergenerational Occupational Mobility in Great Britain and the United States since 1850: Comment. American Economic Review 103(5): 2003–2020.

The M-index in some more detail

- The M-index (mutual information) measures the amount of information shared between A and B.
  - Learning A teaches me something on B.
  - Difference between what I know about B after learning A (a posteriori) and what I knew before (a priori).
- In social mobility context:
  - If effects of social origin are strong, social origin will tell me a lot about the social destination of a person
  - A priori information: based on the (unconditional) marginal distribution of the classes of destination P_Y
  - A posteriori information: based on the conditional distribution of the classes of destination, given the class of origin $P_{y|x}$
- Entropy can be used as a measure for the amount of information: Higher entropy means lower amount of information
  - A priory entropy: $T(P_Y) = -\sum_{k=1}^{K} p(y_k) \ln(p(y_k))$
  - A posteriori entropy: $T(P_Y|X) = -\sum_{j=1}^{J} \sum_{k=1}^{K} p(y_k|x_j) \ln(p(y_k|x_j))$

$$M = T(P_Y) - T(P_Y|X) = \sum_{j=1}^{J} \sum_{k=1}^{K} p(y_k, x_j) \left(\ln p(y_k|x_j) - \ln p(y_k)\right)$$